



Office de la Propriété

Intellectuelle
du Canada

Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office

An agency of
Industry Canada

CA 2372332 A1 2000/11/02

(21) **2 372 332**

**(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION**

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2000/04/17

(51) Cl.Int.⁷/Int.Cl.⁷ C12N 15/29, A01H 1/00, C12N 15/82,
C12N 15/66, C12N 15/52, C12N 9/04

(87) Date publication PCT/PCT Publication Date: 2000/11/02

(71) Demandeur/Applicant:
BASF AKTIENGESELLSCHAFT, DE

(85) Entrée phase nationale/National Entry: 2001/10/29

(72) Inventeurs/Inventors:

(86) N° demande PCT/PCT Application No.: EP 2000/003465

HERBERS, KARIN, DE;
REINDL, ANDREAS, DE;
SCHWENDER, JORG, DE;
LICHTENTHALER, HARTMUT, DE

(87) N° publication PCT/PCT Publication No.: 2000/065036

(74) Agent: ROBIC

(30) Priorité/Priority: 1999/04/27 (199 18 949.8) DE

(54) Titre : SUREXPRESSION D'UNE SEQUENCE D'ADN CODANT POUR UNE 1-DESOXY-D-XYLULOSE-5-
PHOSPHATE REDUCTOISOMERASE A L'INTERIEUR DE PLANTES
(54) Title: OVEREXPRESSION OF A DNA SEQUENCE CODING FOR A 1-DESOXY-D-XYLULOSE-5-PHOSPHATE
REDUCTOISOMERASE IN PLANTS

(57) Abrégé/Abstract:
The invention relates to a method for producing plants containing increased quantities of tocopherols, vitamin K, carotenoids, chlorophylls and polyterpenes by overexpression of a DXPRI gene.



(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG(19) Weltorganisation für geistiges Eigentum
Internationales Büro(43) Internationales Veröffentlichungsdatum
2. November 2000 (02.11.2000)

PCT

(10) Internationale Veröffentlichungsnummer
WO 00/65036 A3

(51) Internationale Patentklassifikation⁷: C12N 15/29, (74) Gemeinsamer Vertreter: BASF AKTIENGESELLSCHAFT; D-67056 Ludwigshafen (DE).

(21) Internationales Aktenzeichen: PCT/EP00/03465 (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) Internationales Anmeldedatum: 17. April 2000 (17.04.2000)

(25) Einreichungssprache: Deutsch (84) Bestimmungsstaaten (regional): ARIPO-Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität: 199 18 949.8 27. April 1999 (27.04.1999) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; D-67056 Ludwigshafen (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): LICHTENTHALER, Hartmut [DE/DE]; Im Kennental 17, D-76227 Karlsruhe (DE). SCHWENDER, Jörg [DE/DE]; Selenbachstrasse 5, D-76327 Pfinztal (DE). REINDL, Andreas [DE/DE]; Albertine-Scherer-Strasse 21, D-67134 Birkenheide (DE). HERBERS, Karin [DE/DE]; Am Hange 6, D-06484 Quedlinburg (DE).

(88) Veröffentlichungsdatum des internationalen Recherchenberichts: 19. April 2001

Veröffentlicht:

— Mit internationalem Recherchenbericht.

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

WO 00/65036 A3

(54) Title: OVEREXPRESSION OF A DNA SEQUENCE CODING FOR A 1-DESOXY-D-XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE IN PLANTS

(54) Bezeichnung: ÜBEREXPRESSION EINER DNA-SEQUENZ CODIEREND FÜR EINE 1-DESOXY-D-XYLULOSE-5-PHOSPHAT REDUKTOISOMERASE IN PFLANZEN

(57) Abstract: The invention relates to a method for producing plants containing increased quantities of tocopherols, vitamin K, carotenoids, chlorophylls and polyterpenes by overexpression of a DXPRI gene.

(57) Zusammenfassung: Verfahren zur Herstellung von Pflanzen mit erhöhtem Gehalt an Tocopherolen, Vitamin K, Carotinoiden, Chlorophyllen und Polyterpenen durch Überexpression eines DXPRI-Gens.

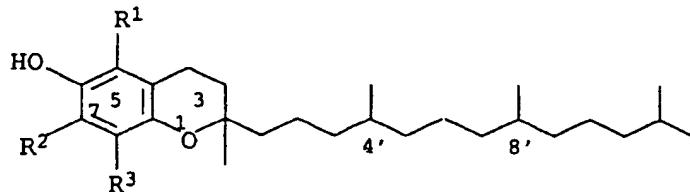
OVEREXPRESSION OF A DNA SEQUENCE CODING FOR A 1-DESOXY-D-
XYLULOSE-5-PHOSPHATE REDUCTOISOMERASE IN PLANTS

The invention relates to a DNA encoding a polypeptide with 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXPRI) activity which originates from plants. In addition, the invention relates to the use of DNA sequences encoding a polypeptide with DXPRI activity which originates from plants for the generation of plants with an elevated tocopherol, carotenoid, vitamin K, chlorophyll and polyterpene content, specifically to the use of

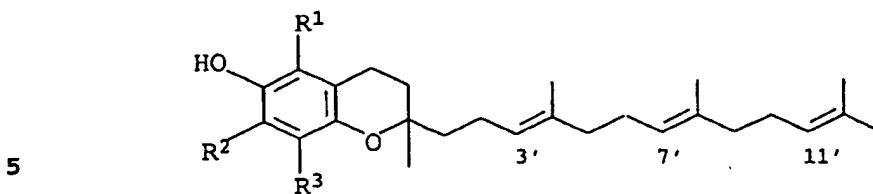
10 the DNA sequence SEQ ID No. 1 or of DNA sequences hybridizing herewith, to a method for the generation of plants with an elevated tocopherol, carotenoid, vitamin K, chlorophyll and polyterpene content, and to the resulting plant itself.

The generation of plants with an elevated sugar, enzyme and amino acid content has hitherto been an important objective in plant molecular genetics. The development of plants with an elevated vitamin content, such as, for example, an elevated tocopherol content, is, however, also of economic interest.

The naturally occurring eight compounds with vitamin E activity are derivatives of 6-chromanol (Ullmann's Encyclopedia of Industrial Chemistry, Vol. A 27 (1996), VCH Verlagsgesellschaft, 20 Chapter 4., 478-488, Vitamin E). The first group (1a-d) is derived from tocopherol, while the second group is composed of tocotrienol derivatives (2a- d):



1a, α -tocopherol: $R^1 = R^2 = R^3 = CH_3$
 1b, β -tocopherol [148-03-8]: $R^1 = R^3 = CH_3$, $R^2 = H$
 1c, γ -tocopherol [54-28-4]: $R^1 = H$, $R^2 = R^3 = CH_3$
 30 1d, δ -tocopherol [119-13-1]: $R^1 = R^2 = H$, $R^3 = CH_3$



2a, α -tocotrienol [1721-51-3]: $R^1 = R^2 = R^3 = CH_3$

2b, β -tocotrienol [490-23-3]: $R^1 = R^3 = CH_3$, $R^2 = H$

10 2c, γ -tocotrienol [14101-61-2]: $R^1 = H$, $R^2 = R^3 = CH_3$

2d, δ -tocotrienol [25612-59-3]: $R^1 = R^2 = H$, $R^3 = CH_3$

α -Tocopherol has great economic importance.

15 The development of crop plants with an elevated tocopherol, vitamin K, carotenoid, chlorophyll and polyterpene content by means of tissue culture or seed mutagenesis and natural selection is set a limit. Thus, on the one hand, it must be possible to appraise, for example, the tocopherol content, or content of the 20 desired catabolite, as early as in the tissue culture stage, and, on the other hand, only those plants whose regeneration from cell cultures to entire plants is successful can be manipulated by tissue culture techniques. Also, crop plants may show undesirable characteristics after mutagenesis and selection, and

25 these characteristics must be reeliminated by, in some cases repeated, back crosses. Also, for example, the increase in tocopherol content would be restricted to crosses between plants of the same species.

30 This is why the genetic engineering approach of isolating essential biosynthesis genes which encode, for example, tocopherol synthesis performance and introducing them into crop plants in a directed fashion is superior to the traditional breeding method. Knowledge of the biosynthesis and its 35 regulation, and identification of genes which affect biosynthesis performance, are prerequisites for this method.

Isoprenoids or terpenoids are composed of a variety of classes of lipid-soluble molecules, and they are formed partially or 40 exclusively from C₅-isoprene units. Pure prenyl lipids (for example carotenoids) are composed of C skeletons based exclusively on isoprene units, while mixed prenyl lipids (for example chlorophylls, tocopherols and vitamin K), have an isoprenoid side chain linked to an aromatic nucleus.

The biosynthesis of prenyl lipids starts with 3 x acetyl-CoA units which are converted into the starting isoprene unit (C₅), namely isopentenyl pyrophosphate (IPP), via β-hydroxymethylglutaryl-CoA (HMG-CoA) and mevalonate. Recent C¹³ 5 in vivo feeding experiments have demonstrated that the IPP formation pathway in various eubacteria, green algae and plant chloroplasts is mevalonate-independent (Figure 1). In this pathway, hydroxyethylthiamine, which is formed by decarboxylation of pyruvate, and glycerolaldehyde-3-phosphate (3-GAP) are first 10 converted into 1-deoxy-D-xylulose-5-phosphate in a "transketolase" reaction mediated by 1-deoxy-D-xylulose-5-phosphate synthase (DOXS) (Lange et al., 1998; Schwender et al., 1997; Arigoni et al., 1997; Lichtenthaler et al., 1997; Sprenger et al., 1997). In an intramolecular 15 rearrangement reaction, this 1-deoxy-D-xylulose-5-phosphate is converted by DXPRI into 2-C-methyl-D-erythritol-4-phosphate and then into IPP (Arigoni et al., 1997; Zeidler et al., 1998). Biochemical data suggest that the mevalonate pathway operates in the cytosol and leads to the formation of phytosterols. The 20 antibiotic mevinolin, a specific mevalonate formation inhibitor, only leads to sterol biosynthesis inhibition in the cytoplasma, while prenyl lipid formation in the plastids remains unaffected (Bach and Lichtenthaler, 1993). In contrast, the mevalonate-independent pathway is located in the plastids and 25 leads predominantly to the formation of carotenoids and plastid prenyl lipids (Schwender et al., 1997; Arigoni et al., 1997).

IPP is in equilibrium with its isomer, dimethylallyl pyrophosphate (DMAPP). Condensation of IPP with DMAPP head to 30 tail results in the monoterpene (C₁₀) geranylpyrophosphate (GPP). Addition of further IPP units results in the sesquiterpene (C₁₅) farnesyl pyrophosphate (FPP), and to the diterpene (C₂₀) geranylgeranyl pyrophosphate (GGPP). Bonding between two GGPP molecules results in the formation of the C₄₀ precursors of 35 carotenoids.

In the case of mixed prenyl lipids, the isoprene side chain, whose length varies, is linked to non-isoprene rings such as, for example, a porphyrine ring in the case of chlorophylls a and b. 40 The chlorophylls and phylloquinones contain a C₂₀ phytyl chain, in which only the first isoprene unit contains a double bond. GGPP is converted by geranylgeranyl pyrophosphate oxidoreductase (GGPPOR) to give phytyl pyrophosphate (PPP), the starting material for the subsequent formation of tocopherols.

The ring structures of the mixed prenyl lipids which lead to the formation of vitamins E and K are quinones whose starting metabolites are derived from the shikimate pathway. The aromatic amino acids phenylalanine or tyrosine are converted into 5 hydroxyphenyl pyruvate, which is dioxygenated to give homogentisic acid. The chorismate is formed, on the one hand, via erythrose-4-phosphate, 3'-dehydroquinate, 3'-dehydroshikimate, shikimate, shikimate-3-phosphate and 5'-enolpyruvylshikimate-3-phosphate (Fig. 1). In this process, 10 fructose-6-phosphate and glycerolaldehyde-3-phosphate are reacted to give xylulose-5-phosphate and erythrose-4-phosphate. The above-described homogentisic acid is subsequently bonded to PPP to form the precursor of α -tocopherol and α -tocopherone, namely 2-methyl-6-phytylquinol. Methylation steps with 15 S-adenosylmethionine as methyl group donor lead first to 2,3-dimethyl-6-phytylquinol, subsequent cyclization leads to γ -tocopherol and further methylation to α -tocopherol (Richter, Biochemie der Pflanzen [Plant biochemistry], Georg Thieme Verlag Stuttgart, 1996).

20 Examples which demonstrate that manipulation of an enzyme may directionally affect metabolite flow can be found in the literature. A direct effect on the quantities of carotenoids in these transgenic tomato plants was measured in experiments on an 25 altered expression of phytoene synthase, which links two GGPP molecules to give 15-cis-phytoene (Fray and Grierson, Plant Mol. Biol. 22(4), 589-602 (1993); Fray et al., Plant J., 8, 693-701 (1995)). As expected, transgenic tobacco plants which have reduced quantities of phenylalanine-ammonium lyase show reduced 30 quantities of phenylpropanoid. The enzyme phenylalanine-ammonium lyase catalyzes the degradation of phenylalanine and thus withdraws it from phenylpropanoid biosynthesis (Bate et al., Proc. Natl. Acad. Sci USA 91 (16): (1994) 7608-7612; Howles et al., Plant Physiol. 112. (1996) 1617-1624).

35 Little is known to date on increasing the metabolite flow for elevating the tocopherol content in plants by overexpression of individual biosynthesis genes. Only WO 97/27285 describes a modification of the tocopherol content by stronger expression or 40 down-regulation of the enzyme p-hydroxyphenylpyruvate dioxygenase (HPPD).

It is an object of the present invention to develop a transgenic plant with an elevated tocopherol, vitamin K, carotenoid, 45 chlorophyll and polyterpene content.

We have found that this object has been achieved by overexpressing a 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXPRI) gene in the plants.

5 To increase the metabolite flow from the primary metabolism into the tocopherol biosynthesis, for example, the formation of 2-C-methyl-D-erythritol-4-P as essential starting substrate for all plastid isoprenoids was increased. To this end, DXPRI activity in transgenic plants was increased by overexpressing the 10 *Arabidopsis thaliana* DXPRI gene. In principle, this can also be achieved by expressing homologous or heterologous DXPRI genes. A nucleotide sequence encoding a DXPRI was described for *E.coli* (Accession Number AB 013300; Kuzuyama et al., 1998; Takahashi et al., 1998).

15

Example 1 describes for the first time a plant DXPRI gene (Fig. 2, SEQ-ID No. 1) from *Arabidopsis thaliana*, which is expressed at a higher level in transgenic plants. To ensure localization in the plastids, a transit signal sequence (Fig. 3, Fig. 4) is 20 arranged upstream of the *Arabidopsis thaliana* DXPRI nucleotide sequence. Fragment A (529 bp) in Figure 4 contains the cauliflower mosaic virus 35S promoter (nucleotides 6909 to 7437 of the cauliflower mosaic virus). Fragment B (259 bp) contains the transketolase transit peptide. Fragment E contains the DXPRI 25 gene. Fragment D (192 bp) contains the polyadenylation signal of gene 3 of the Ti-plasmid pTIACH5 T-DNA (Gielen et al., 1984) to terminate transcription. Another suitable expression cassette is a DNA sequence which encodes a DXPRI gene which hybridizes with a SEQ ID No. 1 and which is derived from other organisms or other 30 plants.

The 2-C-methyl-D-erythritol-4-P, of which greater quantities are now available owing to the additional expression of the DXPRI gene, is reacted further towards tocopherols, carotenoids, 35 vitamin K, chlorophylls and polyterpenes.

The transgenic plants are generated by transforming the plants with a construct comprising the DXPRI gene. Tobacco and oilseed rape were used as model plants for the production of tocopherols, 40 vitamin K, carotenoids, chlorophylls and polyterpenes.

Antisense constructs and homologous or heterologous plant DXPRI genes were transformed independently of one another into plants (Fig. 5). Fragment A (529 bp) in Figure 5 contains the 45 cauliflower mosaic virus 35S promoter (nucleotides 6909 to 7437 of the cauliflower mosaic virus). Fragment B (259 bp) contains the transketolase transit peptide (Figure 3). Fragment E contains

the DXPRI gene in antisense orientation. Fragment D (192 bp) contains the polyadenylation signal of gene 3 of the Ti-plasmid pTIACH5 T-DNA (Gielen et al., 1984) to terminate transcription. Measurements on DXPRI antisense plants showed a drastic decrease 5 in tocopherol and carotenoid contents. This confirms the direct effect of the plastid plant DXPRI on carotenoid and tocopherol synthesis.

The invention relates to the use of the *Arabidopsis thaliana* DNA 10 sequence SEQ ID No. 1 which encodes a DXPRI or its functional equivalents for the generation of a plant with an elevated tocopherol, carotenoid, vitamin K, chlorophyll and polyterpene content. The nucleic acid sequence may be, for example, a DNA or cDNA sequence. Encoding sequences which are suitable for 15 insertion into an expression cassette are, for example, those which encode a DXPRI and which allow the host to overproduce tocopherols, carotenoids, vitamin K, chlorophylls and polyterpenes.

20 The expression cassettes also comprise regulative nucleic acid sequences which govern the expression of the encoding sequence in the host cell. In a preferred embodiment, an expression cassette comprises a promoter upstream, i.e. on the 5'-end of the encoding sequence, and a polyadenylation signal downstream, i.e. on the 25 3'-end, and, if appropriate, further regulatory elements which are linked operatively with the sequence in between which encodes the DXPRI gene. Operative linkage is to be understood as meaning the sequential arrangement of promoter, encoding sequence, terminator and, if appropriate, further regulatory elements in 30 such a way that each of the regulatory elements can fulfil its function as intended when the encoding sequence is expressed. The sequences preferred for operative linkage, but not restricted thereto, are targeting sequences for guaranteeing subcellular 35 localization in the apoplast, in the vacuole, in plastids, in the mitochondrion, in the endoplasmatic reticulum (ER), in the nucleus, in oil bodies or in other compartments, and translation enhancers such as the tobacco mosaic virus 5'-leader sequence (Gallie et al., *Nucl. Acids Res.* 15 (1987), 8693 -8711).

40 As an example, the plant expression cassette can be incorporated into the tobacco transformation vector pBinAR-Hyg. Fig. 6 shows the tobacco transformation vectors pBinAR-Hyg with 35S promoter (A) and pBinAR-Hyg with the seed-specific promoter phaseolin 796 (B):

45

- HPT: hygromycin phosphotransferase
- OCS: octopin synthase terminator

- PNOS: nopaline synthase promoter
- also shown are those restriction cleavage sites which cut the vector only once.

5 A suitable promoter of the expression cassette is, in principle, any promoter which is capable of governing the expression of foreign genes in plants. In particular, a plant promoter or a promoter derived from a plant virus is preferably used.

Particularly preferred is the CaMV 35S promoter from cauliflower 10 mosaic virus (Franck et al., Cell 21 (1980), 285 - 294). As is known, this promoter contains various recognition sequences for transcriptional effectors which in their totality lead to permanent and constitutive expression of the introduced gene (Benfey et al., EMBO J. 8 (1989), 2195-2202).

15

The expression cassette may also comprise a chemically inducible promoter which allows expression of the exogenous DXPRI gene in the plant to be governed at a particular point in time. Such promoters which can be used are, inter alia, for example the PRP1 20 promoter (Ward et al., Plant. Mol. Biol. 22 (1993), 361-366), a salicylic-acid-inducible promoter (WO 95/19443), a benzenesulfonamide-inducible promoter (EP-A 388186), a tetracyclin-inducible promoter (Gatz et al., (1992) Plant J. 2, 397-404), an abscisic-acid-inducible promoter (EP-A 335528) or an 25 ethanol- or cyclohexanone-inducible promoter (WO 93/21334).

Furthermore, particularly preferred promoters are those which ensure expression in tissues or parts of the plant in which, for example, the biosynthesis of tocopherol or its precursors takes 30 place. Promoters which ensure leaf-specific expression must be mentioned in particular. Promoters which must be mentioned are the potato cytosolic FBPase or the potato ST-LSI promoter (Stockhaus et al., EMBO J. 8 (1989), 2445 - 245).

35 A foreign protein was expressed stably in the seeds of transgenic tobacco plants to an extent of 0.67% of the total soluble seed protein with the aid of a seed-specific promoter (Fiedler and Conrad, Bio/Technology 10 (1995), 1090-1094). The expression cassette can therefore contain, for example, a seed-specific 40 promoter (preferably the phaseolin promoter (US 5504200), the USP promoter (Baumlein, H. et al., Mol. Gen. Genet. (1991) 225 (3), 459 - 467) or the LEB4 promoter (Fiedler and Conrad, 1995)), the LEB4 signal peptide, the gene to be expressed and an ER retention signal.

An expression cassette is generated by fusing a suitable promoter with a suitable DXPRI DNA sequence and, preferably, a DNA which is inserted between promoter and DXPRI DNA sequence and which encodes a chloroplast-specific transit peptide, and with a

5 polyadenylation signal, using customary recombination and cloning techniques as they are described, for example, by T. Maniatis, E.F. Fritsch and J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) and by T.J. Silhavy, M.L. Berman and L.W. Enquist,

10 Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and by Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley-Interscience (1987).

15 Particularly preferred sequences are those which ensure targeting into the apoplast, into plastids, into the vacuole, into the mitochondrion or into the endoplasmatic reticulum (ER) or which, due to a lack of suitable operative sequences, ensure that the product remains in the compartment of formation, the cytosol

20 (Kermode, *Crit. Rev. Plant Sci.* 15, 4 (1996), 285-423). Localization in the ER has proved to be particularly advantageous for the amount of protein accumulation in transgenic plants (Schouten et al., *Plant Mol. Biol.* 30 (1996), 781-792).

25 Other expression cassettes which can be used are those whose DNA sequence encodes a DXPRI fusion protein, part of the fusion protein being a transit peptide which governs translocation of the polypeptide. Chloroplast-specific transit peptides which are cleaved off enzymatically from the DXPRI residue after

30 translocation of the DXPRI gene into the chloroplasts are preferred. Particularly preferred is the transit peptide derived from plastid DXPRI or from a functional equivalent of this transit peptide (for example the transit peptide of the Rubisco small subunit or of ferredoxin NADP oxidoreductase).

35 Especially preferred are DNA sequences of three cassettes of the plastid transit peptide of potato plastid transketolase in three reading frames as KpnI/BamHI fragments with an ATG codon in the NcoI cleavage site:

40 pTP09

KpnI_GGTACCATGGCGTCTTCTTCTCTCACTCTCTCAAGCTATCCTCTCGTTCTGTC
CCTCGCCATGGCTCTGCCTCTCTCAACTTTCCCTCTCTCACTTTTCCGGCCTTAA

45 ATCCAATCCAAATATCACCACCTCCCGCCGCGTACTCCTCCTCCGCCGCCGCCGTCG

TAAGGTACCGGCGATTCGTGCCTCAGCTGCAACCGAAACCATAAGAGAAAATGAGACTGCCGGATC
TCC_BamHI

pTP10

5

KpnI_GGTACCATGGCGTCTTCTTCTCACTCTCTCAAGCTATCCTCTCGTTCTGTC
CCTCGCCATGGCTCGCCTCTCTCAACTTTCCCTTCTCACTTTCCGGCCCTAA
ATCCAATCCAATATCACCAACCTCCCGCCGCGTACTCCTCCGCGCCGCCGCCGTG
TAAGGTACCGGCGATTCGTGCCTCAGCTGCAACCGAAACCATAAGAGAAAATGAGACTGCCGCTG

10 GATCC_BamHI

pTP11

KpnI_GGTACCATGGCGTCTTCTTCTCACTCTCTCAAGCTATCCTCTCGTTCTGTC
15 CCTCGCCATGGCTCGCCTCTCTCAACTTTCCCTTCTCACTTTCCGGCCCTAA
ATCCAATCCAATATCACCAACCTCCCGCCGCGTACTCCTCCGCGCCGCCGCCGTG
TAAGGTACCGGCGATTCGTGCCTCAGCTGCAACCGAAACCATAAGAGAAAATGAGACTGCCGGG
ATCC_BamHI

20 The inserted nucleotide sequence encoding a DXPRI can be prepared synthetically, obtained naturally or contain a mixture of synthetic and natural DNA constituents, and may be composed of various heterologous DXPRI gene segments of a variety of organisms. In general, synthetic nucleotide sequences are
25 produced which are equipped with codons which are preferred by plants. These codons which are preferred by plants can be determined from codons with the highest protein frequency expressed in the plant species of the highest interest. When preparing an expression cassette, a variety of DNA fragments may
30 be manipulated in order to obtain a nucleotide sequence which expediently reads in the correct direction and which is equipped with a correct reading frame. Adaptors or linkers may be added to the fragments in order to link the DNA fragments to each other.

35 The promoter and terminator regions may expediently be provided, in the direction of transcription, with a linker or polylinker containing one or more restriction sites for insertion of this sequence. As a rule, the linker has 1 to 10, in most cases 1 to 8, preferably 2 to 6, restriction sites. In general, the linker
40 within the regulatory regions has a size less than 100 bp, frequently less than 60 bp, but at least 5 bp. The promoter may be native, or homologous, or else foreign, or heterologous, to the host plant. The expression cassette comprises, in the 5'-3' direction of transcription, the promoter, a DNA sequence encoding
45 a DXPRI gene, and a region for transcriptional termination.

Various termination regions may be exchanged for each other as desired.

Manipulations which provide suitable restriction cleavage sites 5 or which eliminate the excess DNA or restriction cleavage sites may also be employed. In vitro mutagenesis, primer repair, restriction or ligation may be used in cases where insertions, deletions or substitutions such as, for example, transitions and transversions, are suitable. Complementary ends of the fragments 10 may be provided for ligation in the case of suitable manipulations such as, for example, restriction, chewing back or filling in overhangs for blunt ends.

A procedure which may be of importance for the success according 15 to the invention may be, inter alia, the attachment of the specific ER retention signal SEKDEL (Schouten, A. et al., Plant Mol. Biol. 30 (1996), 781 - 792), thus tripling to quadrupling the average expression level. Other retention signals which occur naturally in plant and animal proteins localized in the ER may 20 also be employed for constructing the cassette.

Preférred polyadenylation signals are plant polyadenylation signals, preferably those which correspond essentially to Agrobacterium tumefaciens T-DNA-polyadenylation signals, in 25 particular those of gene 3 of the T-DNA (octopin synthase) of the Ti-plasmid pTiACH5 (Gielen et al., EMBO J. 3 (1984), 835 et seq.), or functional equivalents.

For example, an expression cassette may comprise a constitutive 30 promoter (preferably the CaMV 35 S promoter), the LeB4 signal peptide, the gene to be expressed, and the ER retention signal. The preferred ER retention signal used is the amino acid sequence KDEL (lysine, aspartic acid, glutamic acid, leucine).

35 The fused expression cassette which encodes a DXPRI gene is preferably cloned into a vector, for example pBin19, which is suitable for transforming Agrobacterium tumefaciens. Agrobacteria transformed with such a vector can then be used in a known manner for transforming plants, in particular crop plants, such as, for 40 example, tobacco plants, for example by bathing wounded leaves or leaf sections in an agrobacterial suspension and subsequently growing them in suitable media. The transformation of plants by agrobacteria is known, inter alia, from F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, 45 Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, pp. 15 - 38. Transgenic plants which comprise integrated into the expression cassette a gene for

expressing a DXPRI gene can be regenerated in a known manner from the transformed cells of the wounded leaves or leaf sections.

To transform a host plant with a DNA encoding a DXPRI, an 5 expression cassette is inserted into a recombinant vector whose vector DNA comprises additional functional regulatory signals, for example sequences for replication or integration. Suitable vectors are described, inter alia, in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), chapter 6/7, pp. 71 - 119 10 (1993).

Using the above-cited recombination and cloning techniques, the expression cassettes can be cloned into suitable vectors which allow their multiplication, for example in *E. coli*. Suitable 15 cloning vectors are, inter alia, pBR322, pUC series, M13mp series and pACYC184. Especially suitable are binary vectors which are capable of replication in *E. coli* and in agrobacteria.

The invention furthermore relates to the use of an expression 20 cassette comprising the DNA sequence SEQ ID No. 1, or DNA sequences hybridizing herewith, for transforming plants, plant cells, plant tissues or parts of plants. The preferred object of the use is an elevated tocopherol, vitamin K, carotenoid, chlorophyll and polyterpene content of the plant.

25 Depending on the choice of promoter, expression may take place specifically in the leaves, in the seeds or in other parts of the plant. Such transgenic plants, their propagation material and the cells, tissues or parts of such plants are a further subject of 30 the present invention.

In addition, the expression cassette may also be employed for transforming bacteria, cyanobacteria, yeasts, filamentous fungi and algae for the purpose of increasing the tocopherol, vitamin 35 K, carotenoid, chlorophyll and polyterpene content.

The transfer of foreign genes into the genome of a plant is termed transformation. It exploits the above-described methods of transforming and regenerating plants from plant tissues or plant 40 cells for transient or stable transformation. Suitable methods are protoplast transformation by polyethylene-glycol-induced DNA uptake, the biolistic method using the gene gun - the so-called particle bombardment method, electroporation, incubation of dry embryos in DNA-containing solution, microinjection and 45 agrobacterium-mediated gene transfer. The abovementioned methods are described in, for example, B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and

Utilization, edited by S.D. Kung and R. Wu, Academic Press (1993), 128 - 143, and in Potrykus, Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991), 205 - 225). The construct to be expressed is preferably cloned into a vector which is suitable 5 for the transformation of *Agrobacterium tumefaciens*, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711).

Agrobacteria transformed with an expression cassette can equally be used in a known manner for transforming plants, in particular 10 crop plants such as cereals, corn, oats, soya, rice, cotton, sugar beet, canola, sunflower, flax, hemp, potato, tobacco, tomato, oilseed rape, alfalfa, lettuce and the various tree, nut and grapevine species, for example by bathing wounded leaves or 15 leaf sections in an agrobacterial suspension and subsequently growing them in suitable media.

Functionally equivalent sequences which encode a DXPRI gene are those sequences which still have the desired functions, despite a differing nucleotide sequence. Functional equivalents thus 20 encompass naturally occurring variants of the sequences described herein, and synthetic nucleotide sequences, for example those obtained by chemical synthesis and adapted to suit the codon usage of a plant.

25 Functional equivalents are also to be understood as meaning, in particular, natural or artificial mutations of an originally isolated sequence which encodes a DXPRI which continue to show the desired function. Mutations encompass substitutions, additions, deletions, exchanges or insertions of one or more 30 nucleotide residues. Thus, the present invention also encompasses, for example, those nucleotide sequences which are obtained by modifying the DXPRI nucleotide sequence. The purpose of such a modification may be, for example, the further limitation of the encoding sequence contained therein or else, 35 for example, the insertion of further restriction enzyme cleavage sites.

Functional equivalents are also those variants whose function is attenuated or increased compared with the starting gene, or gene 40 fragment.

Also suitable are artificial DNA sequences as long as they mediate the desired characteristic, for example an elevated tocopherol content in the plant, by overexpression of the DXPRI 45 gene in crop plants, as described above. Such artificial DNA sequences can be identified, for example, by back translation of proteins with DXPRI activity which have been constructed by means

13

of molecular modeling, or else by *in vitro* selection. Especially suitable are encoding DNA sequences which have been obtained by back translating a polypeptide sequence in accordance with the host-plant-specific codon usage. An expert skilled in the art of 5 plant genetic methods will readily be able to identify the specific codon usage by computer evaluations of other known genes of the plant to be transformed.

Further suitable equivalent nucleic acid sequences which must be 10 mentioned are sequences which encode fusion proteins, a DXPRI polypeptide or a functionally equivalent portion of these being a constituent of the fusion protein. The second part of the fusion protein may be, for example, another enzymatically active polypeptide, or an antigenic polypeptide sequence with the aid of 15 which detection of DXPRI expression is possible (for example myc-tag or his-tag). However, it is preferably a regulatory protein sequence such as, for example, a signal or transit peptide which leads the DXPRI protein to the desired site of action.

20

An elevated tocopherol, vitamin K, chlorophyll, carotenoid and polyterpene content is to be understood as meaning for the purposes of the present invention the artificially acquired capability of an increased biosynthetic performance regarding 25 these compounds by functional overexpression of the DXPRI gene in the plant in comparison with the non-genetically-modified plant for at least one plant generation.

The tocopherol biosynthesis site, for example, is generally the 30 leaf tissue, so that leaf-specific expression of the DXPRI gene is meaningful. However, it is obvious that tocopherol biosynthesis need not be limited to the leaf tissue but may also take place in a tissue-specific fashion in the other remaining parts of the plant, for example in fatty seeds.

35

The constitutive expression of the exogenous DXPRI gene is also advantageous. On the other hand, inducible expression may also be desirable.

40 The expression efficacy of the transgenically expressed DXPRI gene can be determined, for example, *in vitro* by shoot meristem propagation. In addition, altered expression of the DXPRI gene with regard to type and level, and its effect on tocopherol biosynthesis performance may be tested on test plants in 45 greenhouse experiments.

14

The invention furthermore relates to transgenic plants, transformed with an expression cassette comprising the sequence SEQ ID No. 1 or DNA sequences hybridizing herewith, and to transgenic cells, tissues, parts and propagation material of such 5 plants. Especially preferred are transgenic crop plants such as, for example, barley, wheat, rye, corn, oats, soya, rice, cotton, sugar beet, canola, sunflower, flax, hemp, potato, tobacco, tomato, oilseed rape, alfalfa, lettuce and the various tree, nut and grapevine species.

10

Plants for the purposes of the invention are mono- and dicotyledonous plants, or algae.

Since this biosynthetic pathway is a metabolic pathway which is 15 exclusively located in the chloroplasts, it offers optimal target enzymes for the development of inhibitors. Since, according to current knowledge, no enzyme which is identical with, or similar to, *Arabidopsis thaliana* DXPRI is present in other higher 20 organisms, it can be assumed that inhibitors should have a very specific action on plants. The site of action of an inhibitor, namely fosmidomycin (3-(N-formyl-N-hydroxyamino)propylphosphonic acid; Fujisawa Pharmaceutical Co.) was identified as being a DXPRI. The biochemical assay shows efficient inhibition of the enzymatic activity (Fig. 7). The following abbreviations were 25 used in Fig. 7: DOX = 1-deoxy-D-xylulose, ME = methylerythritol. The same action is found in a plant assay in which barley seedlings are subjected to fosmidomycin treatment and then examined for their chlorophyll and carotenoid content. Both substances, which are derived from precursors of the isoprenoid 30 metabolism, are greatly reduced in terms of quantity (Fig. 8).

Overexpression of the DXPRI-encoding gene sequence SEQ ID No. 1 and SEQ ID No. 3 in a plant allows, in principle, an improved resistance to DXPRI inhibitors to be achieved. Transgenic plants 35 generated thus are also subject-matter of the invention.

Other subject-matters of the invention are:

- Methods of transforming a plant, which comprise introducing, 40 into a plant cell, into callus tissue, an entire plant or plant protoplasts, expression cassettes comprising a DNA sequence SEQ ID No. 1 or DNA sequences hybridizing herewith.
- The use of the expression cassette comprising a DNA sequence 45 SEQ ID No. 1 or DNA sequences hybridizing herewith for generating plants with an elevated resistance to DXPRI

15

inhibitors by stronger expression of the DNA sequence SEQ ID No. 1, or DNA sequences hybridizing herewith.

5 - The use of the DNA sequence SEQ ID No. 1 or DNA sequences hybridizing herewith for the generation of plants with an elevated tocopherol, vitamin K, chlorophyll, carotenoid and polyterpene content by expressing a DXPRI DNA sequence in plants.

10 The invention is now illustrated by the examples which follow, but not limited thereto:

General cloning methods

15 The cloning steps carried out within the scope of the present invention, e.g. restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids onto nitrocellulose and nylon membranes, linking DNA fragments, transformation of *E. coli* cells, growing bacteria, 20 multiplying phages and sequence analysis of recombinant DNA, were carried out as described by Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6.

The bacterial strains used hereinbelow (*E. coli*, XL-I Blue) were 25 obtained from Stratagene. The agrobacterial strain used for the transformation of plants (*Agrobacterium tumefaciens*, C58C1 with plasmid pGV2260 or pGV3850kan) was described by Deblaere et al. in Nucl. Acids Res. 13 (1985), 4777. Alternatively, the agrobacterial strain LBA4404 (Clontech) or other suitable strains 30 may also be employed. Vectors which can be used for cloning are the vectors pUC19 (Yanish-Perron, Gene 33 (1985), 103 - 119) pBluescript SK- (Stratagene), pGEM-T (Promega), pZero (Invitrogen), pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984), 8711 - 8720) and pBinAR (Höfgen and Willmitzer, Plant Science 66 35 (1990), 221 - 230).

Sequence analysis of recombinant DNA

Recombinant DNA molecules were sequenced using a laser 40 fluorescence DNA sequencer by Licor (available from MWG Biotech, Ebersbach) following the method of Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA 74 (1977), 5463 - 5467).

Example 1

Cloning of *Arabidopsis thaliana* DXPRI

5 Starting from the sequence of the *E. coli* DXPRI, which had been deposited in the gene library, other bacterial protein sequences which were homologous with DXPRI were identified in gene databases. A comparison of the protein sequences, each of which only had a length of 400 amino acids, showed several conserved 10 amino acid sequence motives. Such a motive showed homologies with a deposited genomic *Arabidopsis* sequence (Accession Number AB009053).

Since the bacterial DXPRI sequences show a conserved amino acid 15 sequence close to the putative N-terminus, the beginning of a functional section of the *Arabidopsis* DXPRI sequence was localized very exactly in deposited genomic sequences. The C-terminal end of the sequence (stop codon) was found by comparison with the EST clone (Accession Number AA586087). A 1215 20 bp DXPRI fragment was cloned and studied for enzymatic functionality by means of heterologous expression.

mRNA was isolated from *Arabidopsis thaliana* (var. Columbia), and cDNA was prepared (following the instructions of the manufacturer 25 Stratgene). PCR primers were derived from sequences AB009053 and AA586087, and with their aid a 1215 bp DNA fragment from the cDNA which had been prepared was amplified. The primer ATRv3 has a BamHI cleavage site and is chosen in such a way that, following restriction digest and ligation into pBluescript or pET5b 30 (expression plasmid; Promega), the encoding sequence starting from the N-terminal first conserved sequence is ligated into the reading frame of the protein translation.

Atrv3 5' TCAGGGATCCGGCGCCTCGTCAATCT 3'
35 Attr1 5' GACGAATTCTTTCTCCAACAAACCAATTCT 3'

The primers Atrv3 and Attr1 contained a BamHI and an EcoRI cleavage site, respectively (in each case underlined). The PCR product (Atrv3/Attr1) was purified by means of the Gene-Clean-Kit 40 (Dianova GmbH, Hilden) and digested with BamHI and EcoRI. For the ligation, vector pET5b was also cleaved with BamHI and EcoRI. The ligation products were transformed into *E. coli* XL1Blue (Stratagene).

45 Plasmid pET5bAtr contains a gene fragment encoding *Arabidopsis thaliana* DXPRI. Its sequence was determined (Figure 2, SEQ-ID No. 1). The nucleotide sequence obtained from plasmid pET5bAtr can be

compared with sequences AB009053 and AA586087. Accordingly, the genomic sequence AB009053 contains 10 introns.

Example 2

5

Cloning of the *Arabidopsis thaliana* DXPRI into the expression vector pET5bAtr, and detection of the enzymatic activity

The expression vector pET5b (Promega) is an expression vector for 10 the expression of recombinant proteins in *E. coli*. The plasmid is derived from pBR322 and carries a bacteriophage T7 promoter for expression. For expression, the plasmid is multiplied in an *E. coli* strain which carries an inducible gene for T7 polymerase (for example JM109(DE3); Promega). Expression of the recombinant 15 protein is activated by inducing T7 polymerase.

pET5bAtr encodes a fusion protein which is 420 amino acids long. Amino acids 1 to 14 are derived from pET5b (fusion peptide; Figure 3). Amino acids 15 to 420 are derived from the cloned 20 DXPRI fragment (Fig. 2). In Fig. 3, the DNA sequence for the fusion peptide is underlined. Based on the entire sequence, the molecular weight for the protein is calculated as 45.6 kD.

The transgenic strain was incubated in the growth medium "2x YT" 25 (per 1 l: Bacto-tryptone 16 g, yeast extract 10 g, NaCl 5g). The cells were grown at 37°C to an OD_{560 nm} of 0.6. After IPTG (1 mM) had been added, they were cultured for a further 10 minutes at 37°C and then for a further 4 hours at 22° until they were harvested. The cells were centrifuged off and washed in 1% NaCl. 30 After the cells had been disrupted (50 to 500 ml of cell culture (OD_{560 nm} of 1.0) by means of a French press, a crude protein extract was used for enzyme tests (in 4 ml of extraction buffer (Tris/HCl (pH 7.5) 100 mM, MgCl₂ 5 mM, DTT 2 mM, PMSF 0.1 mM). For storage, the crude extracts together with 20% of glycerol were 35 frozen at -20°C.

For the enzyme test, 15 µl (diluted to 1 to 7 mg of protein/ml) of protein extract were incubated for 30 minutes at 30°C with MnCl₂ (1 mM), NaF (5 mM), NADPH₂ (0.5 mM) and 40 ¹⁴C-1-deoxy-D-xylulose-5-phosphate (0.25 mM, 3 kBq). The reaction was stopped at 100°C (30 seconds) by adding CaCl₂ (to 100 mM) and alkaline phosphatase (0.5 units). The product was dephosphorylated for 2 hours at 30°C. Detection was carried out by subjecting the product to separation by thin-layer chromatography 45 on silica gel 60 (Merck) with acetone/ethyl acetate/water (50+50+2) with subsequent evaluation by Instant Imager. This gives 1-deoxy-D-xylulose (DOX; Rf 0.4) and methylerythritol (ME;

18

Rf 0.2). Figure 9 shows the evaluation by thin-layer chromatography and autoradiography after heterologous expression of the *Arabidopsis thaliana* DXPRI in *E. coli* and enzyme assay with the use of various total protein concentrations (μ g protein/ μ l). K = control *E. coli* JM 109 (DE3) with plasmid pET5b without DXPRI. Samples: *E. coli* JM 109 (DE3) with plasmid pET5b with *Arabidopsis thaliana* DXPRI. The formation of ME can be inhibited efficiently by using fosmidomycin at various concentrations. This shows the mode of action of fosmidomycin as DXPRI inhibitor.

Example 3

Preparation of the substrate 1-deoxy-D-xylulose-5-phosphate (DOXP) for the enzyme assay

To prepare DOXP, DOXS cloned from *Chlamydomonas reinhardtii* was used (pET5b, *E. coli* JM109(DE3)).

Enzyme extracts of IPTG-induced *E. coli* cells were incubated with [$3-^{14}\text{C}$]-pyruvate and DL-GAP. After 30 minutes, the reaction was stopped by heat-denaturation of the proteins. After centrifugation, the conversion of the radioactive pyruvate was checked by means of TLC/autoradiography, and the supernatant was used as substrate for the reductoisomerase.

DOXP as reaction product was identified with reference to the following criteria:

1. This gives rise to a radioactive product which, after treatment with alkaline phosphatase, behaves in a less polar fashion in TLC separations. This suggests that a phosphorylated product of ^{14}C -pyruvate and GAP was formed.
2. In TLC (silica gel, acetone/ethyl acetate/water 50/50/2), the dephosphorylated product migrates in the same way as a synthetic 1-deoxy-D-xylulose sample.

The restriction batch contained protein extract (20 μ l/100 μ l of batch), Tris/HCl (pH 7.5) 100 mM, DTT 2 mM, MgCl₂ 5 mM, Na-EDTA 500 μ M, PMSF 100 μ M, NaF 5 mM, TPP 1 mM, sodium pyruvate 1 mM, sodium [$2-^{14}\text{C}$]-pyruvate 20 kBq/100 μ l, and DL glycerol aldehyde-3-phosphate 3.75 mM.

Example 4

Cloning of *Arabidopsis thaliana* DXPRI into the plant transformation vector pBin19AR-TP

5

To clone DXPRI into a binary vector, the primers were chosen in such a way that, following restriction digest and ligation into pBin19AR-TP (Promega), the encoding sequence starting from the N-terminal first conserved sequence is ligated into the reading 10 frame of the protein translation.

AtrvpBin1 5' TCAGGATCCGGCGCCTCGTCAATCT 3'
AttrpBin2 5' GACCCCGGGTCTTCCAACAACCAATTCT 3'

15 The primers AtrvpBin1 and AttrpBin2 contained a BamHI or SmaI cleavage site, respectively (in each case underlined). The PCR product (AtrvpBin1/AttrpBin2) was purified by means of the Gene-Clean-Kit (Dianova GmbH, Hilden) and digested with BamHI and SmaI. For the ligation, the vector pBin19AR-TP was also cleaved 20 with BamHI and SmaI, and additionally contained the potato transketolase transit peptide downstream of the CaMV 35S promoter. The transit peptide ensures localization in the plastid. The construct is shown in Figure 4.

25 Example 5

Generation of DXPRI antisense constructs

The following primers were chosen to clone DXPRI into a binary 30 vector in antisense orientation.

AtrvpBin3 5' TCACCCGGGGGCCTCGTCAATCT 3'
AttrpBin4 5' GACGGATCCTCTTCCAACAACCAATTCT 3'

35 The primers AtrvpBin3 and AttrpBin4 contained an SmaI or BamHI cleavage site, respectively (in each case underlined). The PCR product (AtrvpBin3/AttrpBin4) was purified by means of the Gene-Clean-Kit (Dianova GmbH, Hilden) and digested with SmaI and BamHI. For the ligation, the vector pBin19AR-TP was also cleaved 40 with SmaI and BamHI, and additionally contained the potato transketolase transit peptide downstream of the CaMV 35S promoter. The transit peptide ensures localization in the plastid. The construct is shown in Figure 5.

20

Tobacco plants with a reduced DXPRI activity were selfed and the resulting seed was harvested. To analyze the plants further, seeds from the F1 generation were used.

5 All antisense plants studied showed clear differences with regard to plant size. Plants were found which had the same size as the wild type, but also very small plants. The subsequent generations were, therefore, not uniform. This also applies to the reduction in DXPRI activity, which was not uniform within one line, i.e. a 10 line cannot be defined by a specific reduction in DXPRI activity, but the lines segregate (a comparable phenomenon is shown by sedoheptulose-1,7-bisphosphatase antisense tobacco plants; cf. Harrison et al. 1998, *Planta* 204: 27-36).

15 Biomass analysis showed a correlation between reduction in DXPRI activity and biomass reduction.

Example 6

20 Extraction and detection of tocopherol

Extraction method:

- Leaf material, 100 mg fresh weight
- 25 - Extraction buffer: 80% ethanol, 10 mM Hepes pH 7.0, 1 mM ascorbate
- Extraction: 1:5 (w/v)
- Incubation for 30 minutes at 50°C
- No centrifugation
- 30 - Addition of 1/2 volume of n-hexane to the extract
- Vortexing and centrifugation (5 minutes, room temperature)
- Recovery of the deep green top phase
- Repetition of the n-hexane extraction with the bottom phase
- Combination of the n-hexane phases
- 35 - Drying in vacuum (2-3 hours per ml of n-hexane at room temperature)
- Redissolution of the residue in approx. 1/5 of the original n-hexane volume
- Application of 30 - 50 µl onto HPLC
- 40 - HPLC detection of tocopherol
- Detection of fluorescence: excitation at 295 nm, emission at 330 nm
- Column: RP-18 (Nucleosil 100, C18, 3µm, Knauer)
- Isocratic system: n-hexane plus 0.2% of 2-propanol
- 45 - Flow rate: 0.8 ml/min (pressure: 110 bar)
- Standards by Sigma or Merck

21

- Chromatography time: 15 minutes

Example 7

5 Extraction of phenolic substances from leaves, and HPLC analysis

The extraction of phenolic substances from leaves was carried out as described by Yao et al., The Plant Cell, 7 (1995), 1787.

10 Example 8

Generation of transgenic tobacco plants

(*Nicotiana tabacum* L. cv. Samsun NN).

15

To generate transgenic tobacco plants with an altered prenyl lipid content, tobacco leaf disks were transformed with DXPRI sequences (SEQ ID No. 1) cloned into the transformation vector pBin19AR-TP, as described in Example 4. To transform the tobacco 20 plants, 10 ml of an overnight culture of *Agrobacterium tumefaciens* which had grown under selection conditions was centrifuged, the supernatant was discarded, and the bacteria were resuspended in an equal volume of antibiotic-free medium. Leaf disks of sterile plants (approx. diameter 1 cm) were bathed in 25 this bacterial suspension in a sterile Petri dish. The leaf disks were subsequently plated in Petri dishes on MS medium (Murashige and Skoog, Physiol. Plant (1962) 15, 473) supplemented with 2% sucrose and 0.8% Bacto agar. After incubation in the dark at 25°C for two days, they were transferred to MS medium supplemented 30 with 100 mg/l kanamycin, 500 mg/l claforan, 1 mg/l benzylaminopurin (BAP), 0.2 mg/l naphthylacetic acid (NAA), 1.6% glucose and 0.8% Bacto agar, and grown on (16 hours light/8 hours dark). Growing shoots were transferred to hormone-free MS medium 35 supplemented with 2% sucrose, 250 mg/l claforan and 0.8% Bacto agar.

Example 9

Generation of transgenic oilseed rape plants (*Brassica napus*)

40

The generation of the transgenic oilseed rape plants which have an altered prenyl lipid content followed in principle a procedure described by Bade, J.B. and Damm, B. (in Gene Transfer to Plants, Potrykus, I. and Spangenberg, G., eds, Springer Lab Manual, 45 Springer Verlag, 1995, 30-38), which also gives the compositions of the media and buffers used.

The transformations were performed with *Agrobacterium tumefaciens* strain LBA4404 (Clontech GmbH, Heidelberg). The binary vectors used were the binary constructs which have already been described in Example 4 which comprised the *Arabidopsis thaliana* DXPRI total 5 cDNA (SEQ ID No. 1). In all the binary vectors used here, the NOS terminator sequence was replaced by the gene 3 polyadenylation signal of Ti-plasmid pTIACH5 T-DNA (Gielen et al., 1984) for transcription termination. *Brassica napus* seeds were surface-sterilized with 70% (v/v) ethanol, washed for 10 minutes 10 in H₂O at 55°C, incubated for 20 minutes in 1% strength hypochlorite solution (25% v/v Teepol, 0.1% v/v Tween 20) and washed for 20 minutes six times with sterile H₂O. The seeds were dried on filter paper for three days and 10-15 seeds were germinated in a glass flask containing 15 ml of germination 15 medium. The roots and apices were removed from several seedlings (approx. length 10 cm), and the remaining hypocotyls were cut into sections approx. 6 mm in length. The approx. 600 explants thus obtained were washed for 30 minutes in 50 ml of basal medium and transferred into a 300 ml flask. After 100 ml of callus 20 induction medium had been added, the cultures were incubated for 24 hours at 100 rpm.

An overnight culture of the agrobacterial strain was established in Luria broth medium supplemented with kanamycin (20 mg/l) at 25 29°C, and 2 ml of this were incubated in 50 ml of Luria broth medium without kanamycin for 4 hours at 29°C to an OD₆₀₀ of 0.4 - 0.5. After the culture had been pelleted for 25 minutes at 2000 rpm, the cell pellet was resuspended in 25 ml of basal medium. The bacterial concentration in the solution was brought to an 30 OD₆₀₀ of 0.3 by adding more basal medium.

The callus induction medium was removed from the oilseed rape explants using sterile pipettes, 50 ml of agrobacterial suspension were added, the cultures were mixed carefully and 35 incubated for 20 minutes. The agrobacterial suspension was removed, the oilseed rape explants were washed for 1 minute with 50 ml of callus induction medium, and 100 ml of callus induction medium were subsequently added. Cocultivation was performed for 24 hours on an orbital shaker at 100 rpm. Cocultivation was 40 stopped by removing callus induction medium, and the explants were washed twice with 25 ml of wash medium for 1 minute each time and twice for 60 minutes with 100 ml of wash medium each time, at 100 rpm. The wash medium together with the explants was transferred into 15 cm Petri dishes, and the medium was removed 45 using sterile pipettes.

23

For regeneration, batches of 20-30 explants were transferred into 90 mm Petri dishes containing 25 ml of shoot induction medium supplemented with kanamycin. The Petri dishes were sealed with two layers of Leukopor and incubated at 25°C and 2000 lux at 5 photoperiods of 16 hours light/8 hours dark. Every 12 days, the developing calli were transferred to fresh Petri dishes containing shoot induction medium. All further steps for regenerating entire plants were carried out as described by Bade, J.B. and Damm, B. (in Gene Transfer to Plants, Potrykus, I. and 10 Spangenberg, G., eds, Springer Lab Manual, Springer Verlag, 1995, 30-38).

Example 10

15 Increasing the tocopherol biosynthesis in tobacco

The overexpression of *Arabidopsis thaliana* DXPRI in tobacco was carried out as described in Example 8.

20 Tobacco plants which had been transformed with suitable constructs were grown in the greenhouse. The α -tocopherol content of the entire plant and of the seeds of the plant was subsequently determined. The α -tocopherol concentration was increased in all cases relative to the untransformed plant.

25

Example 11

Increasing the tocopherol biosynthesis in oilseed rape

30 The *Arabidopsis thaliana* DXPRI cDNA (SEQ-No. 1) was provided with a CaMV35S promoter and overexpressed in oilseed rape using the 35S promoter. In parallel, the seed-specific phaseoline gene promoter was used to increase the tocopherol content specifically 35 in the rape seed. Oil seed rape plants which had been transformed with suitable constructs were grown in the greenhouse. The α -tocopherol content of the entire plant and of the seeds of the plant was subsequently determined. The α -tocopherol concentration was increased in all cases relative to the untransformed plant.

40

45

We claim:

1. The use of DNA sequences encoding a 5 1-deoxy-D-xylulose-5-phosphate reductoisomerase for generating plants with an elevated tocopherol, vitamin K, carotenoid, chlorophyll and polyterpene content.
2. The use of a DNA sequence SEQ ID No. 1 or of a DNA sequence 10 hybridizing herewith, encoding a 1-deoxy-D-xylulose-5-phosphate reductoisomerase for the generation of plants with an elevated tocopherol, vitamin K, carotenoid, chlorophyll and content.
- 15 3. A polyterpene method for the generation of plants with an elevated tocopherol, vitamin K, carotenoid, chlorophyll and polyterpene content, which comprises expressing, in plants, a DNA sequence SEQ ID No. 1 or a DNA sequence hybridizing herewith.
- 20 4. A method of transforming a plant, which comprises introducing, into a plant cell, into callus tissue, an entire plant or plant cell protoplasts, an expression cassette comprising a promoter and a DNA sequence SEQ ID No. 1 or a DNA sequence hybridizing herewith.
- 25 5. A method of transforming plants as claimed in claim 4, wherein transformation is effected with the aid of the strain *Agrobacterium tumefaciens*, electroporation or the particle 30 bombardment method.
6. A plant with an elevated tocopherol, vitamin K, carotenoid, chlorophyll and polyterpene content comprising an expression cassette as set forth in claim 4.
- 35 7. A plant as claimed in claim 6 selected from the group consisting of soya, canola, barley, oats, wheat, oilseed rape, corn or sunflower.
- 40 8. A plant as claimed in claim 7 selected from the group consisting of soya, canola, barley, oats, wheat, oilseed rape, corn or sunflower.

SEQUENZPROTOKOLL

<110> BASF AG

<120> Ueberexpression einer DNA-Sequenz codierend fuer eine 1-Desoxy-D-Xylulose-5-Phosphat Reduktoisomerase in Pflanzen

<130> 990380

<140> 0050-49943

<141> 1999-04-27

<160> 4

<170> PatentIn Vers. 2.0

<210> 1

<211> 1221

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1221)

<400> 1

gcg	cct	cgt	caa	tct	tgg	gat	gga	cca	aaa	ccc	atc	tct	atc	gtt	gga	48
Ala	Pro	Arg	Gln	Ser	Trp	Asp	Gly	Pro	Lys	Pro	Ile	Ser	Ile	Val	Gly	
1	5							10				15				

tct	act	ggt	tct	att	ggc	act	cag	aca	ttg	gat	att	gtg	gct	gag	aat	96
Ser	Thr	Gly	Ser	Ile	Gly	Thr	Gln	Thr	Leu	Asp	Ile	Val	Ala	Glu	Asn	
20								25				30				

cct	gac	aaa	ttc	aga	gtt	gtg	gct	cta	gct	gct	ggg	tcg	aat	gtt	act	144
Pro	Asp	Lys	Phe	Arg	Val	Val	Ala	Leu	Ala	Ala	Gly	Ser	Asn	Val	Thr	
35								40				45				

cta	ctt	gct	gat	cag	gta	agg	aga	ttt	aag	cct	gca	ttg	gtt	gct	gtt	192
Leu	Leu	Ala	Asp	Gln	Val	Arg	Arg	Phe	Lys	Pro	Ala	Leu	Val	Ala	Val	
50								55				60				

aga	aac	gag	tca	ctg	att	aat	gag	ctt	aaa	gag	gct	tta	gct	gat	ttg	240
Arg	Asn	Glu	Ser	Leu	Ile	Asn	Glu	Leu	Lys	Glu	Ala	Leu	Ala	Asp	Leu	
65								70				75		80		

gac	tat	aaa	ctc	gag	att	att	cca	gga	gag	caa	gga	gtg	att	gag	gtt	288
Asp	Tyr	Lys	Leu	Glu	Ile	Ile	Pro	Gly	Glu	Gln	Gly	Val	Ile	Glu	Val	

WO 00/65036

PCT/EP00/03465

2

85

90

95

gcc cga cat ccc gaa gct gta acc gtt acc gga ata gta ggt tgt 336
 Ala Arg His Pro Glu Ala Val Thr Val Val Thr Gly Ile Val Gly Cys
 100 105 110

gcg gga cta aag cct acg gtt gct gca att gaa gca gga aag gac att 384
 Ala Gly Leu Lys Pro Thr Val Ala Ala Ile Glu Ala Gly Lys Asp Ile
 115 120 125

gct ctt gca aac aaa gag aca tta atc gca ggt ggt cct ttc gtg ctt 432
 Ala Leu Ala Asn Lys Glu Thr Leu Ile Ala Gly Gly Pro Phe Val Leu
 130 135 140

ccg ctt gcc aac aaa cat aat gta aag att ctt ccg gca gat tca gaa 480
 Pro Leu Ala Asn Lys His Asn Val Lys Ile Leu Pro Ala Asp Ser Glu
 145 150 155 160

cat tct gcc ata ttt cag tgt att caa ggt ttg cct gaa ggc gct ctg 528
 His Ser Ala Ile Phe Gln Cys Ile Gln Gly Leu Pro Glu Gly Ala Leu
 165 170 175

cgc aag ata atc ttg act gca tct ggt gga gct ttt agg gat tgg cct 576
 Arg Lys Ile Ile Leu Thr Ala Ser Gly Gly Ala Phe Arg Asp Trp Pro
 180 185 190

gtc gaa aag cta aag gaa gtt aaa gta gcg gat gcg ttg aag cat cca 624
 Val Glu Lys Leu Lys Glu Val Lys Val Ala Asp Ala Leu Lys His Pro
 195 200 205

aac tgg aac atg gga aag aaa atc act gtg gac tct gct acg ctt ttc 672
 Asn Trp Asn Met Gly Lys Ile Thr Val Asp Ser Ala Thr Leu Phe
 210 215 220

aac aag ggt ctt gag gtc att gaa gcg cat tat ttg ttt gga gct gag 720
 Asn Lys Gly Leu Glu Val Ile Glu Ala His Tyr Leu Phe Gly Ala Glu
 225 230 235 240

tat gac gat ata gag att gtc att cat ccg caa agt atc ata cat tcc 768
 Tyr Asp Asp Ile Glu Ile Val Ile His Pro Gln Ser Ile Ile His Ser
 245 250 255

atg att gaa aca cag gat tca tct gtg ctt gct caa ttg ggt tgg cct 816
 Met Ile Glu Thr Gln Asp Ser Ser Val Leu Ala Gln Leu Gly Trp Pro
 260 265 270

gat atg cgt tta ccg att ctc tac acc atg tca tgg ccc gat aga gtt 864
 Asp Met Arg Leu Pro Ile Leu Tyr Thr Met Ser Trp Pro Asp Arg Val
 275 280 285

cct tgt tct gaa gta act tgg cca aga ctt gac ctt tgc aaa ctc ggt	912																																																																																							
Pro Cys Ser Glu Val Thr Trp Pro Arg Leu Asp Leu Cys Lys Leu Gly																																																																																								
290	295	300		tca ttg act ttc aag aaa cca gac aat gtg aaa tac cca tcc atg gat	960	Ser Leu Thr Phe Lys Lys Pro Asp Asn Val Lys Tyr Pro Ser Met Asp		305	310	315	320	cct gct tat gct gct gga cga gct gga ggc aca atg act gga gtt ctc	1008	Leu Ala Tyr Ala Ala Gly Arg Ala Gly Gly Thr Met Thr Gly Val Leu		325	330	335		agc gcc gcc aat gag aaa gct gtt gaa atg ttc att gat gaa aag ata	1056	Ser Ala Ala Asn Glu Lys Ala Val Glu Met Phe Ile Asp Glu Lys Ile		340	345	350		agc tat ttg gat atc ttc aag gtt gtg gaa tta aca tgc gat aaa cat	1104	Ser Tyr Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His		355	360	365		cga aac gag ttg gta aca tca ccg tct ctt gaa gag att gtt cac tat	1152	Arg Asn Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr		370	375	380		gac ttg tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt	1200	Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly		385	390	395	400	gct agg cca gtt cat gca tga	1221	Ala Arg Pro Val His Ala		405		<210> 2		<211> 406		<212> PRT		<213> <i>Arabidopsis thaliana</i>		<400> 2		Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly			1	5	10	15	Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn			20	25	30		Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr			35	40	45	
300																																																																																								
tca ttg act ttc aag aaa cca gac aat gtg aaa tac cca tcc atg gat	960																																																																																							
Ser Leu Thr Phe Lys Lys Pro Asp Asn Val Lys Tyr Pro Ser Met Asp																																																																																								
305	310	315	320	cct gct tat gct gct gga cga gct gga ggc aca atg act gga gtt ctc	1008	Leu Ala Tyr Ala Ala Gly Arg Ala Gly Gly Thr Met Thr Gly Val Leu		325	330	335		agc gcc gcc aat gag aaa gct gtt gaa atg ttc att gat gaa aag ata	1056	Ser Ala Ala Asn Glu Lys Ala Val Glu Met Phe Ile Asp Glu Lys Ile		340	345	350		agc tat ttg gat atc ttc aag gtt gtg gaa tta aca tgc gat aaa cat	1104	Ser Tyr Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His		355	360	365		cga aac gag ttg gta aca tca ccg tct ctt gaa gag att gtt cac tat	1152	Arg Asn Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr		370	375	380		gac ttg tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt	1200	Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly		385	390	395	400	gct agg cca gtt cat gca tga	1221	Ala Arg Pro Val His Ala		405		<210> 2		<211> 406		<212> PRT		<213> <i>Arabidopsis thaliana</i>		<400> 2		Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly			1	5	10	15	Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn			20	25	30		Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr			35	40	45									
315	320																																																																																							
cct gct tat gct gct gga cga gct gga ggc aca atg act gga gtt ctc	1008																																																																																							
Leu Ala Tyr Ala Ala Gly Arg Ala Gly Gly Thr Met Thr Gly Val Leu																																																																																								
325	330	335		agc gcc gcc aat gag aaa gct gtt gaa atg ttc att gat gaa aag ata	1056	Ser Ala Ala Asn Glu Lys Ala Val Glu Met Phe Ile Asp Glu Lys Ile		340	345	350		agc tat ttg gat atc ttc aag gtt gtg gaa tta aca tgc gat aaa cat	1104	Ser Tyr Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His		355	360	365		cga aac gag ttg gta aca tca ccg tct ctt gaa gag att gtt cac tat	1152	Arg Asn Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr		370	375	380		gac ttg tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt	1200	Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly		385	390	395	400	gct agg cca gtt cat gca tga	1221	Ala Arg Pro Val His Ala		405		<210> 2		<211> 406		<212> PRT		<213> <i>Arabidopsis thaliana</i>		<400> 2		Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly			1	5	10	15	Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn			20	25	30		Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr			35	40	45																	
335																																																																																								
agc gcc gcc aat gag aaa gct gtt gaa atg ttc att gat gaa aag ata	1056																																																																																							
Ser Ala Ala Asn Glu Lys Ala Val Glu Met Phe Ile Asp Glu Lys Ile																																																																																								
340	345	350		agc tat ttg gat atc ttc aag gtt gtg gaa tta aca tgc gat aaa cat	1104	Ser Tyr Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His		355	360	365		cga aac gag ttg gta aca tca ccg tct ctt gaa gag att gtt cac tat	1152	Arg Asn Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr		370	375	380		gac ttg tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt	1200	Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly		385	390	395	400	gct agg cca gtt cat gca tga	1221	Ala Arg Pro Val His Ala		405		<210> 2		<211> 406		<212> PRT		<213> <i>Arabidopsis thaliana</i>		<400> 2		Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly			1	5	10	15	Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn			20	25	30		Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr			35	40	45																									
350																																																																																								
agc tat ttg gat atc ttc aag gtt gtg gaa tta aca tgc gat aaa cat	1104																																																																																							
Ser Tyr Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His																																																																																								
355	360	365		cga aac gag ttg gta aca tca ccg tct ctt gaa gag att gtt cac tat	1152	Arg Asn Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr		370	375	380		gac ttg tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt	1200	Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly		385	390	395	400	gct agg cca gtt cat gca tga	1221	Ala Arg Pro Val His Ala		405		<210> 2		<211> 406		<212> PRT		<213> <i>Arabidopsis thaliana</i>		<400> 2		Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly			1	5	10	15	Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn			20	25	30		Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr			35	40	45																																	
365																																																																																								
cga aac gag ttg gta aca tca ccg tct ctt gaa gag att gtt cac tat	1152																																																																																							
Arg Asn Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr																																																																																								
370	375	380		gac ttg tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt	1200	Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly		385	390	395	400	gct agg cca gtt cat gca tga	1221	Ala Arg Pro Val His Ala		405		<210> 2		<211> 406		<212> PRT		<213> <i>Arabidopsis thaliana</i>		<400> 2		Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly			1	5	10	15	Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn			20	25	30		Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr			35	40	45																																									
380																																																																																								
gac ttg tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt	1200																																																																																							
Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly																																																																																								
385	390	395	400	gct agg cca gtt cat gca tga	1221	Ala Arg Pro Val His Ala		405		<210> 2		<211> 406		<212> PRT		<213> <i>Arabidopsis thaliana</i>		<400> 2		Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly			1	5	10	15	Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn			20	25	30		Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr			35	40	45																																																	
395	400																																																																																							
gct agg cca gtt cat gca tga	1221																																																																																							
Ala Arg Pro Val His Ala																																																																																								
405																																																																																								
<210> 2																																																																																								
<211> 406																																																																																								
<212> PRT																																																																																								
<213> <i>Arabidopsis thaliana</i>																																																																																								
<400> 2																																																																																								
Ala Pro Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly																																																																																								
1	5	10	15																																																																																					
Ser Thr Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn																																																																																								
20	25	30																																																																																						
Pro Asp Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr																																																																																								
35	40	45																																																																																						

WO 00/65036

PCT/EP00/03465

4

Leu Leu Ala Asp Gln Val Arg Arg Phe Lys Pro Ala Leu Val Ala Val
 50 55 60

Arg Asn Glu Ser Leu Ile Asn Glu Leu Lys Glu Ala Leu Ala Asp Leu
 65 70 75 80

Asp Tyr Lys Leu Glu Ile Ile Pro Gly Glu Gln Gly Val Ile Glu Val
 85 90 95

Ala Arg His Pro Glu Ala Val Thr Val Val Thr Gly Ile Val Gly Cys
 100 105 110

Ala Gly Leu Lys Pro Thr Val Ala Ala Ile Glu Ala Gly Lys Asp Ile
 115 120 125

Ala Leu Ala Asn Lys Glu Thr Leu Ile Ala Gly Gly Pro Phe Val Leu
 130 135 140

Pro Leu Ala Asn Lys His Asn Val Lys Ile Leu Pro Ala Asp Ser Glu
 145 150 155 160

His Ser Ala Ile Phe Gln Cys Ile Gln Gly Leu Pro Glu Gly Ala Leu
 165 170 175

Arg Lys Ile Ile Leu Thr Ala Ser Gly Gly Ala Phe Arg Asp Trp Pro
 180 185 190

Val Glu Lys Leu Lys Glu Val Lys Val Ala Asp Ala Leu Lys His Pro
 195 200 205

Asn Trp Asn Met Gly Lys Lys Ile Thr Val Asp Ser Ala Thr Leu Phe
 210 215 220

Asn Lys Gly Leu Glu Val Ile Glu Ala His Tyr Leu Phe Gly Ala Glu
 225 230 235 240

Tyr Asp Asp Ile Glu Ile Val Ile His Pro Gln Ser Ile Ile His Ser
 245 250 255

Met Ile Glu Thr Gln Asp Ser Ser Val Leu Ala Gln Leu Gly Trp Pro
 260 265 270

Asp Met Arg Leu Pro Ile Leu Tyr Thr Met Ser Trp Pro Asp Arg Val
 275 280 285

Pro Cys Ser Glu Val Thr Trp Pro Arg Leu Asp Leu Cys Lys Leu Gly
 290 295 300

Ser Leu Thr Phe Lys Lys Pro Asp Asn Val Lys Tyr Pro Ser Met Asp

WO 00/65036

PCT/EP00/03465

5

305	310	315	320
-----	-----	-----	-----

Leu Ala Tyr Ala Ala Gly Arg Ala Gly Gly Thr Met Thr Gly Val Leu			
325	330	335	

Ser Ala Ala Asn Glu Lys Ala Val Glu Met Phe Ile Asp Glu Lys Ile			
340	345	350	

Ser Tyr Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His			
355	360	365	

Arg Asn Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr			
370	375	380	

Asp Leu Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly			
385	390	395	400

Ala Arg Pro Val His Ala			
405			

<210> 3

<211> 1289

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(1263)

<400> 3

atg gct agc atg act ggt gga cag caa atg ggt cgg gat ccg gcg cct	48
Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Asp Pro Ala Pro	
1 5 10 15	

cgt caa tct tgg gat gga cca aaa ccc atc tct atc gtt gga tct act	96
Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly Ser Thr	
20 25 30	

ggt tct att ggc act cag aca ttg gat att gtg gct gag aat cct gac	144
Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn Pro Asp	
35 40 45	

aaa ttc aga gtt gtg gct cta gct gct ggt tcg aat gtt act cta ctt	192
Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr Leu Leu	
50 55 60	

gct gat cag gta agg aga ttt aag cct gca ttg gtt gct gtt aga aac	240
Ala Asp Gln Val Arg Arg Phe Lys Pro Ala Leu Val Ala Val Arg Asn	

WO 00/65036

PCT/EP00/03465

6

65

70

75

80

gag tca ctg att aat gag ctt aaa gag gct tta gct gat ttg gac tat 288
 Glu Ser Leu Ile Asn Glu Leu Lys Glu Ala Leu Ala Asp Leu Asp Tyr
 85 90 95

aaa ctc gag att att cca gga gag caa gga gtg att gag gtt gcc cga 336
 Lys Leu Glu Ile Ile Pro Gly Glu Gln Gly Val Ile Glu Val Ala Arg
 100 105 110

cat ccc gaa gct gta acc gtt acc gga ata gta ggt tgt gcg gga 384
 His Pro Glu Ala Val Thr Val Val Thr Gly Ile Val Gly Cys Ala Gly
 115 120 125

cta aag cct acg gtt gct gca att gaa gca gga aag gac att gct ctt 432
 Leu Lys Pro Thr Val Ala Ala Ile Glu Ala Gly Lys Asp Ile Ala Leu
 130 135 140

gca aac aaa gag aca tta atc gca ggt ggt cct ttc gtg ctt ccg ctt 480
 Ala Asn Lys Glu Thr Leu Ile Ala Gly Gly Pro Phe Val Leu Pro Leu
 145 150 155 160

gcc aac aaa cat aat gta aag att ctt ccg gca gat tca gaa cat tct 528
 Ala Asn Lys His Asn Val Lys Ile Leu Pro Ala Asp Ser Glu His Ser
 165 170 175

gcc ata ttt cag tgt att caa ggt ttg cct gaa ggc gct ctg cgc aag 576
 Ala Ile Phe Gln Cys Ile Gln Gly Leu Pro Glu Gly Ala Leu Arg Lys
 180 185 190

ata atc ttg act gca tct ggt gga gct ttt agg gat tgg cct gtc gaa 624
 Ile Ile Leu Thr Ala Ser Gly Gly Ala Phe Arg Asp Trp Pro Val Glu
 195 200 205

aag cta aag gaa gtt aaa gta gcg gat gcg ttg aag cat cca aac tgg 672
 Lys Leu Lys Glu Val Lys Val Ala Asp Ala Leu Lys His Pro Asn Trp
 210 215 220

aac atg gga aag aaa atc act gtg gac tct gct acg ctt ttc aac aag 720
 Asn Met Gly Lys Lys Ile Thr Val Asp Ser Ala Thr Leu Phe Asn Lys
 225 230 235 240

ggt ctt gag gtc att gaa gcg cat tat ttg ttt gga gct gag tat gac 768
 Gly Leu Glu Val Ile Glu Ala His Tyr Leu Phe Gly Ala Glu Tyr Asp
 245 250 255

gat ata gag att gtc att cat ccg caa agt atc ata cat tcc atg att 816
 Asp Ile Glu Ile Val Ile His Pro Gln Ser Ile Ile His Ser Met Ile
 260 265 270

WO 00/65036

PCT/EP00/03465

7

gaa aca cag gat tca tct gtg ctt gct caa ttg ggt tgg cct gat atg . 864
 Glu Thr Gln Asp Ser Ser Val Ala Gln Leu Gly Trp Pro Asp Met
 275 280 285

cgt tta ccg att ctc tac acc atg tca tgg ccc gat aga gtt cct tgt 912
 Arg Leu Pro Ile Leu Tyr Thr Met Ser Trp Pro Asp Arg Val Pro Cys
 290 295 300

tct gaa gta act tgg cca aga ctt gac ctt tgc aaa ctc ggt tca ttg 960
 Ser Glu Val Thr Trp Pro Arg Leu Asp Leu Cys Lys Leu Gly Ser Leu
 305 310 315 320

act ttc aag aaa cca gac aat gtg aaa tac cca tcc atg gat ctt gct 1008
 Thr Phe Lys Lys Pro Asp Asn Val Lys Tyr Pro Ser Met Asp Leu Ala
 325 330 335

tat gct gct gga cga gct gga ggc aca atg act gga gtt ctc agc gcc 1056
 Tyr Ala Ala Gly Arg Ala Gly Gly Thr Met Thr Gly Val Leu Ser Ala
 340 345 350

gcc aat gag aaa gct gtt gaa atg ttc att gat gaa aag ata agc tat 1104
 Ala Asn Glu Lys Ala Val Glu Met Phe Ile Asp Glu Lys Ile Ser Tyr
 355 360 365

ttg gat atc ttc aag gtt gtg gaa tta aca tgc gat aaa cat cga aac 1152
 Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His Arg Asn
 370 375 380

gag ttg gta aca tca ccg tct ctt gaa gag att gtt cac tat gac ttg 1200
 Glu Leu Val Thr Ser Pro Ser Leu Glu Ile Val His Tyr Asp Leu
 385 390 395 400

tgg gca cgt gaa tat gcc gcg aat gtg cag ctt tct tct ggt gct agg 1248
 Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly Ala Arg
 405 410 415

cca gtt cat gca tga agaattggtt gttggaaagaa gaattc 1289
 Pro Val His Ala
 420

<210> 4
 <211> 420
 <212> PRT
 <213> Arabidopsis thaliana

<400> 4
 Met Ala Ser Met Thr Gly Gly Gln Gln Met Gly Arg Asp Pro Ala Pro

WO 00/65036

PCT/EP00/03465

8

1

5

10

15

Arg Gln Ser Trp Asp Gly Pro Lys Pro Ile Ser Ile Val Gly Ser Thr
 20 25 30

Gly Ser Ile Gly Thr Gln Thr Leu Asp Ile Val Ala Glu Asn Pro Asp
 35 40 45

Lys Phe Arg Val Val Ala Leu Ala Ala Gly Ser Asn Val Thr Leu Leu
 50 55 60

Ala Asp Gln Val Arg Arg Phe Lys Pro Ala Leu Val Ala Val Arg Asn
 65 70 75 80

Glu Ser Leu Ile Asn Glu Leu Lys Glu Ala Leu Ala Asp Leu Asp Tyr
 85 90 95

Lys Leu Glu Ile Ile Pro Gly Glu Gln Gly Val Ile Glu Val Ala Arg
 100 105 110

His Pro Glu Ala Val Thr Val Val Thr Gly Ile Val Gly Cys Ala Gly
 115 120 125

Leu Lys Pro Thr Val Ala Ala Ile Glu Ala Gly Lys Asp Ile Ala Leu
 130 135 140

Ala Asn Lys Glu Thr Leu Ile Ala Gly Gly Pro Phe Val Leu Pro Leu
 145 150 155 160

Ala Asn Lys His Asn Val Lys Ile Leu Pro Ala Asp Ser Glu His Ser
 165 170 175

Ala Ile Phe Gln Cys Ile Gln Gly Leu Pro Glu Gly Ala Leu Arg Lys
 180 185 190

Ile Ile Leu Thr Ala Ser Gly Gly Ala Phe Arg Asp Trp Pro Val Glu
 195 200 205

Lys Leu Lys Glu Val Lys Val Ala Asp Ala Leu Lys His Pro Asn Trp
 210 215 220

Asn Met Gly Lys Lys Ile Thr Val Asp Ser Ala Thr Leu Phe Asn Lys
 225 230 235 240

Gly Leu Glu Val Ile Glu Ala His Tyr Leu Phe Gly Ala Glu Tyr Asp
 245 250 255

Asp Ile Glu Ile Val Ile His Pro Gln Ser Ile Ile His Ser Met Ile
 260 265 270

Glu Thr Gln Asp Ser Ser Val Leu Ala Gln Leu Gly Trp Pro Asp Met
275 280 285

Arg Leu Pro Ile Leu Tyr Thr Met Ser Trp Pro Asp Arg Val Pro Cys
290 295 300

Ser Glu Val Thr Trp Pro Arg Leu Asp Leu Cys Lys Leu Gly Ser Leu
305 310 315 320

Thr Phe Lys Lys Pro Asp Asn Val Lys Tyr Pro Ser Met Asp Leu Ala
325 330 335

Tyr Ala Ala Gly Arg Ala Gly Gly Thr Met Thr Gly Val Leu Ser Ala
340 345 350

Ala Asn Glu Lys Ala Val Glu Met Phe Ile Asp Glu Lys Ile Ser Tyr
355 360 365

Leu Asp Ile Phe Lys Val Val Glu Leu Thr Cys Asp Lys His Arg Asn
370 375 380

Glu Leu Val Thr Ser Pro Ser Leu Glu Glu Ile Val His Tyr Asp Leu
385 390 395 400

Trp Ala Arg Glu Tyr Ala Ala Asn Val Gln Leu Ser Ser Gly Ala Arg
405 410 415

Pro Val His Ala
420